REMARKS

Claims 1-17 were pending. Claims 1, 6-7, and 9-12 are amended herewith. Support for these amendments is found throughout the specification at, *e.g.*, page 8, lines 1-3. The amendments of claims 9-12 correct minor typographical errors. Therefore, it is believed that no new matter is added. Claims 1-17 are currently pending. No claim has been allowed.

Formal Matters

Applicants gratefully acknowledge the receipt and entry of the communications dated March 28, 2002 and June 28, 2002 and the awarding of the priority date of March 8, 2001 in view of Japanese Application Number 2001-065799.

Rejection Under 35 U.S.C. § 112, Second Paragraph

Claims 3, 4, 6, 7, 11, and 12 are rejected under 35 U.S.C. § 112, second paragraph as allegedly being indefinite. According to the Examiner, claim 3 is indefinite in the recitation of "*Escherichia coli*" in view of the comma after the recitation. The Examiner asserts that the recitation of "10,000 to 100,000" does not clearly indicate whether these are daltons or kilodaltons. The Examiner also asserts that claim 7 is indefinite in its recitation of the acronym "ATP". Applicants respectfully traverse these rejections.

Applicants submit that the comma after the recitation "*Escherichia coli*" fails to render the claim indefinite. *Escherichia coli* or *E. coli* is not a thermophilic bacteria. The claim includes the use of cell extracts from one of three different sources: (1) *E. coli*, (2) thermophilic bacteria, and (3) yeast. The specification clearly discloses thermophilic bacteria as a separate source of cell extracts and provides an example of this class in its disclosure of *Thermus thermophilus*. *See* the specification at page 5, lines 7-18. Therefore, the claim language is not indefinite.

Claim 6 is amended herewith to clarify that the recitation of 10,000 daltons to 100,000 daltons. Thus, the rejection is moot.

While the ordinary artisan would fully understand the meaning of ATP as used in claim 7, the claim is amended herewith to replace "ATP" with "adenosine triphosphate" in an effort to expedite prosecution of the instant application, rendering the instant rejection moot.

In view of the above, Applicants respectfully submit that the basis for these rejections may be removed.

Rejection Under 35 U.S.C. § 103 (a)

Claims 1-17 are rejected under 35 U.S.C. § 103 (a) as allegedly being unpatentable over Spirin et al., *Science* 242:1162-64 (1998) taken with Hendrickson, *Science* 254:51-58 (1991) and Japanese Patent Application No. A-2000-17569. The Examiner argues that Spirin teaches a similar cell-free protein synthesis system using a creatine phosphokinase as an ATP regenerating system and the use of prokaryotic and eukaryotic cell extracts. The Examiner acknowledges that Spirin lacks any teaching regarding cell-free protein synthesis that incorporate heavy atom-labeled amino acids or use dialysis methods. According to the Examiner, new developments in biotechnology based on preparative cell-free translation systems continuous action are anticipated. The Examiner points to Hendrickson's disclosure of the application of MAD for the structure determination of protein structure using heavy atoms in proteins, asserting that one of ordinary skill in the art would be motivated to combine Hendrickson with Spirin. The Examiner also argues that the cited Japanese application teaches a synthesis system using a dialysis system. Applicants traverse this rejection.

The cited combination of references lacks the necessary motivation to modify the continuous flow system of Spirin to produce a protein suitable for x-ray crystallography using heavy atom amino acids in a dialysis system. Spirin expressly discloses the advantage of the described cell free protein synthesis system is the continuous flow action in the system. *See, e.g.*, Spirin at page 1163 and Figure 2 (comparing the continuous flow system to a non-flowing system). Any advances anticipated by Spirin relate to the use of the continuous flow cell-free protein synthesis system, not just to any cell-free system. *See id.* Thus, to combine the disclosure of Spirin with that of the Japanese application to modify the continuous flow system of Spirin to that of a non-flow system would completely change the principle of operation of Spirin's system. *See* MPEP § 2143.02 ("If the proposed modification or combination of the prior art would change the principle of operation of the prior art invention being modified, then the teachings of the references are not

sufficient to render the claims prima facie obvious." (citations omitted)). Therefore, Spirin cannot be properly combined with the Japanese application.

In the absence of the Japanese application, the combination of Spirin and Hendrickson fail to teach each and every element of the claimed methods. Spirin lacks any disclosure regarding the use of dialysis in a cell-free protein synthesis. In fact, Spirin expressly teaches away from using a cell-free protein synthesis system that does not use continuous flow action. Hendrickson fails to remedy this deficiency. Hendrickson reviews the use of wavelength anomalous diffraction (MAD) in various structure determinations including that of proteins containing heavy atoms. However, Hendrickson contains no suggestion or teaching as to how make such proteins, and cell-free protein synthesis systems are not addressed. Thus, the combination of Spirin and Hendrickson fail to render the claimed methods *prima facie* obvious.

The combination of the Japanese application and Hendrickson also fails to render the claimed methods *prima facie* obvious because neither reference teaches or suggests that a cell-free system can be used to generate heavy atom-containing proteins. While the Japanese application appears to disclose a cell-free system using dialysis, there is no disclosure whatsoever regarding the use of such a system to generate proteins suitable for x-ray crystallography. The abstract of the cited Japanese application also lacks any disclosure regarding, *e.g.*, the source of the cell extract and the use of an ATP regenerating system, and the usefulness of these features in generating proteins with heavy atom amino acids. Hendrickson fails to remedy these deficiencies. Therefore, the combination of the Japanese application and Hendrickson do not render the claimed methods *prima facie* obvious.

Furthermore, Applicants submit that if there was some motivation to combine the teachings of these references to produce the heavy atom-containing proteins for x-ray crystallographic analysis more easily in an *in vitro* system, this was entirely missed by practitioners. Both Hendrickson and Spirin were published more than 10 years before the priority date of the instant application. Thus, while the usefulness of heavy atom-containing proteins for MAD analysis was known, the ordinary artisan did not use any of the cell-free protein synthesis systems available to generate such proteins. This is likely to due to the common belief that cell-free protein synthesis systems are inherently unstable, even fragile, and thus not a good system for producing a protein

with a consistently high heavy-atom incorporation rate. *See, e.g.*, Madlin et al. at Abstract (stating that it is a common belief that cell-free translation systems are "inherently unstable, even fragile") (Exhibit A). In point of fact, the ordinary artisan employed an *in vivo* system in met- *E. coli* to produce heavy atom-containing proteins using, *e.g.*, selenomethionine. *See* Hendrickson et al. (Exhibit B). Numerous complications clouded the potential success of the *in vivo* systems being employed. *See* Exhibit B at page 1670. For example, in Hendrickson, the authors list solubility, oxidation, sufficient heavy atom replacement rates, and temperature sensitivity as problems for selenomethionyl proteins. *Id.* This combination of unpredictability in the known systems for heavy atom-containing proteins with the inherent instability of the cell-free protein synthesis system *teaches away* from using the claimed method to produce such proteins. *See* MPEP § 2145 (X)(1)(3) ("The totality of the prior art must be considered, and proceeding contrary to accepted wisdom in the art is evidence of nonobviousness." (citations omitted)). Thus, while the employment of the cell-free protein synthesis system as described by Applicants to make heavy atom-containing proteins is a relatively simple invention, it is obvious only in hindsight based on Applicants' disclosure. A rigorous examination of the references and the accepted wisdom in the field fail to provide a motivation to combine these references or a reasonable expectation of success in such a combination. Accordingly, the cited references fail to render the claimed methods *prima facie* obvious.

Finally, the cited combination of references fail to render the product of the claimed methods obvious. Neither the Japanese application nor Spirin teach or suggest the use of the disclosed systems to generate proteins with heavy atom amino acids. Hendrickson is completely silent with regards to methods of making proteins with heavy atom amino acids, and therefore contains no suggestion on the properties or characteristics of the resulting proteins. As none of the references suggest or even address the nature of a protein containing at least 80% substitution of a heavy atom amino acid, making an assertion that such a protein is obvious in view of the cited references is without scientific basis.

In view of the above, Applicants respectfully submit that the basis for these rejections may be removed.

CONCLUSION

In view of the above, each of the presently pending claims in this application is believed to be in immediate condition for allowance. Accordingly, the Examiner is respectfully requested to withdraw the outstanding rejection of the claims and to pass this application to issue. If it is determined that a telephone conference would expedite the prosecution of this application, the Examiner is invited to telephone the undersigned at the number given below.

In the event the U.S. Patent and Trademark office determines that an extension and/or other relief is required, applicant petitions for any required relief including extensions of time and authorizes the Commissioner to charge the cost of such petitions and/or other fees due in connection with the filing of this document to Deposit Account No. 03-1952 referencing docket no. 251002009300. However, the Commissioner is not authorized to charge the cost of the issue fee to the Deposit Account.

Dated: June 3, 2004

Respectfully submitted,

By _____
Laurie L. Hill, Ph.D.
Registration No.: 51,804
MORRISON & FOERSTER LLP
3811 Valley Centre Drive, Suite 500
San Diego, California 92130
(858) 720-7955

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: Plants apparently contain a suicide system directed at ribosomes

Kairat Madin, Tatsuya Sawasaki, Tomio Ogasawara, and Yaeta Endo*

Department of Applied Chemistry, Faculty of Engineering, Ehime University, Matsuyama 790-8577, Japan

Edited by Harry F. Noller, University of California, Santa Cruz, CA, and approved November 15, 1999 (received for review August 17, 1999)

Current cell-free protein synthesis systems can synthesize proteins with high speed and accuracy, but produce only a low yield because of their instability over time. Here we describe the preparation of a highly efficient but also robust cell-free system from wheat embryos. We first investigated the source of the instability of existing systems in light of endogenous ribosome-inactivating proteins and found that ribosome inactivation by tritin occurs already during extract preparation and continues during incubation for protein synthesis. Therefore, we prepared our system from extensively washed embryos that are devoid of contamination by endosperm, the source of tritin and possibly other inhibitors. In a batch system, we observed continuous translation for 4 h, and sucrose density gradient analysis showed formation of large polysomes, indicating high protein synthesis activity. When the reaction was performed in a dialysis bag, enabling the continuous supply of substrates together with the continuous removal of small byproducts, translation proceeded for >60 h, yielding 1–4 mg of enzymatically active proteins, and 0.6 mg of a 126-kDa tobacco mosaic virus protein, per milliliter of reaction volume. Our results demonstrate that plants contain endogenous inhibitors of translation and that after their elimination the translational apparatus is very stable. This contrasts with the common belief that cell-free translation systems are inherently unstable, even fragile. Our method is useful for the preparation of large amounts of active protein as well as for the study of protein synthesis itself.

The development of a system capable of synthesizing any desired protein on a preparative scale is one of the most important endeavors in biotechnology today. Three strategies are currently being used: chemical synthesis, *in vivo* expression, and cell-free protein synthesis. The first two methods have severe limitations: chemical synthesis is not feasible for the synthesis of long peptides because of low yield, and *in vivo* expression can produce only those proteins that do not affect the physiology of the host cell (1–3). Cell-free translation systems, in contrast, can synthesize proteins with high speed and accuracy, approaching *in vivo* rates (4–5), and they can express proteins that would interfere with cell physiology. However, they are relatively inefficient because of their instability (6).

Because cell free systems nonetheless have great potential for large scale protein synthesis, many efforts have been made to increase their efficiency. Spirin *et al.* (7) proposed a continuous flow cell-free translation system, in which a solution containing amino acids and energy sources is supplied to the reaction chamber through a filtration membrane. This design is significantly more efficient than conventional batch systems: The reaction works for tens of hours and produces hundreds of micrograms per milliliter of reaction volume (7–9). Recently, several modified versions of the Spirin system have been reported (10–13). Kigawa *et al.* showed that, by using a dialysis membrane to facilitate the continuous supply of substrates and removal of byproducts, an *Escherichia coli*-coupled transcription–translation system yields as much as 6 mg of protein per milliliter of reaction volume (12). This high productivity can, however, only be expected with fairly small proteins such as Ras protein (21 kDa) or chloramphenicol acetyltransferase (26 kDa). The problem with larger proteins is that with the increasing molecular weight of the mRNAs their degradation by endogenous *E. coli* ribonuclease(s) also increases. Kawarasaki *et al.* showed that in a wheat germ cell-free system translational efficiency increases after neutralization of endogenous ribonucleases and phosphatases with copper ions and antiphosphatase antibodies (13). For their improvements, these groups focused on modifying the reaction chamber and/or optimizing the reaction conditions while using conventional extracts. We used a different approach, instead focusing on clarifying the nature of the instability of the extracts.

We concentrated on wheat germ cell-free systems because they have numerous advantages such as low cost, easy availability in large amounts, low endogenous incorporation, and the capacity to synthesize high-molecular-weight proteins. Moreover they are eukaryotic systems and hence more suitable for the expression of eukaryotic proteins. After we discovered that the mechanism of action of the ricin toxin is ribosome inactivation (14–16), many other ribosome-inactivating proteins (RIPs) with identical mechanism of action have been found in higher plants (17). Most commonly these toxins are single-chain proteins, and they inhibit protein synthesis by removing a single adenine residue in a universally conserved stem-loop structure of 28S ribosomal RNA (14–17). Although the biological function of the RIPs is not known, it is generally believed that they are important for cell defense (17). The most widely studied example is an antiviral effect during infection by several plant viruses (18). As originally proposed by Ready *et al.* (19), the explanation for the antiviral activity of RIPs is that, when a cell wall is damaged, the RIP is released into the cytosol, where it inactivates ribosomes, thereby preventing virus replication. Tritin, found in wheat seeds and thought to be localized mainly in the endosperm, is such a single-chain RIP (20). Initially, it was reported that wheat embryonic ribosomes are resistant to this protein (20–22), which would render any contamination with tritin inconsequential.

To improve protein synthesis in wheat germ cell-free systems, we started with the hypothesis that the embryonic ribosomes are in fact susceptible to tritin. In this case, contamination of wheat germ preparations with tritin-containing endosperm fragments would be fatal. Accordingly, we prepared our cell-free system

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: RIP, ribosome-inactivating protein; TMV, tobacco mosaic virus; GFP, green fluorescent protein; DHFR, dihydrofolate reductase.

*To whom reprint requests should be addressed. E-mail: yendo@en3.ehime-u.ac.jp.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

from extensively washed embryos and indeed found that the system became far more active.

In addition to the benefit of a better protein synthesis system, these results shed new light on the translational apparatus itself: Although it is usually seen as a rather fragile apparatus, it appears instead to be very stable: so stable, in fact, that plants seem to have developed a suicide mechanism (the RIPs) directed against the translational apparatus, further emphasizing its crucial role in cell physiology. We believe that the strategy we followed to improve the wheat cell-free system—elimination of endogenous translational inhibitors—is equally applicable for other systems.

Materials and Methods

General. The following procedures were either described or cited previously (9, 14–15, 23–24): determination of RNA *N*-glycosidase activity, analysis of cell-free protein synthesis, sucrose density gradient analysis of polysomes, determination of proteins, the sources of m7GpppG, ribonucleotide triphosphates, SP6 RNA polymerase, T7 RNA polymerase, human placental ribonuclease inhibitor (133 units/ml), L-[U-^{14}C]leucine, MTX immobilized on agarose, creatine kinase, spermidine, and the 20 amino acids. Dialysis membrane (molecular weight cutoff 12,000–14,000, regenerated cellulose, Viskase Sales, Chicago), the nonionic detergent Nonidet P-40, and proteinase inhibitor E64 were purchased from Nakarai Tesque (Kyoto). The luciferase assay kit (PiccaGene) was from Wako Pure Chemical (Osaka). Low and high molecular weight marker kits (Rainbow marker) were from Amersham Pharmacia. Recombinant forms of luciferase and green fluorescent protein (GFP) (S65T) that were used as standards were purchased from Promega and CLONTECH, respectively. Plasmid pCaMV35S-sGFP(S65T)-NOS3′(25) carrying the GFP gene was kindly provided by Y. Niwa (School of Food and Nutritional Sciences, University of Shizuoka, Japan), and plasmid pSP-Luc$^+$ carrying luciferase was obtained from Promega. Plasmid pTLW3 (26), covering the tobacco mosaic virus (TMV) genome, was a generous gift from Y. Watanabe (University of Tokyo).

Purification of Wheat Embryos and Extract Preparation. Wheat seeds were ground in a mill (Roter Speed Mill model pulverisette 14, Fritsh, Germany), then were sieved through a 710- to 850-mm mesh. Embryos were selected with the solvent flotation method of Erickson and Blobel (27) by using a solvent containing cyclohexane and carbon tetrachloride (240:600, vol/vol). Damaged embryos and contaminants were discarded, and intact embryos were dried overnight in a fume hood. To remove contaminating endosperm, the embryos were washed three times with 10 vol of water under vigorous stirring, and then were sonicated for 3 min in a 0.5% solution of Nonidet P-40 by using a Bronson model 2210 sonicator (Yamato, Japan). Finally, the embryos were washed once more in the sonicator with sterile water.

Preparation of the Cell-Free Extract. The method used is a slight modification of the procedure described by Erickson and Blobel (27). Washed embryos were ground to a fine powder in liquid nitrogen. Five grams of the powder were added to 5 ml of 2 × buffer A (40 mM Hepes, pH 7.6/100 mM potassium acetate/5 mM magnesium acetate/2 mM calcium chloride/4 mM DTT/0.3 mM of each of the 20 amino acids). The mixture was briefly vortexed and then was centrifuged at 30,000 × *g* for 30 min. The resulting supernatant was subjected to gel-filtration on a G-25 (fine) column, equilibrated with two volumes of buffer A. The void volume was collected and centrifuged at 30,000 × *g* for 10 min. The final supernatant was adjusted to 200 A260/ml with buffer A, was divided into small aliquots, and was stored in liquid nitrogen until use.

Cell-Free Translation. In the batch system, 50 μl of reaction mixture contained 12.5 μl of extract (thus 24%); final concentrations of the various ingredients are 24 mM Hepes/KOH (pH 7.8), 1.2 mM ATP, 0.25 mM GTP, 16 mM creatine phosphate, 0.45 mg/ml creatine kinase, 2 mM DTT, 0.4 mM spermidine, 0.3 mM of each of the 20 amino acids including [^{14}C]leucine (2 μCi/ml), 2.5 mM magnesium acetate, 100 mM potassium acetate, 50 μg/ml of deacylated tRNA prepared from wheat embryos, 0.05% Nonidet P-40, 1 μM E-64 as proteinase inhibitor, 0.005% NaN_3, and 7.2 μg (0.02 nmol) of dihydrofolate reductase (DHFR) mRNA. The extract was not treated with micrococcal nuclease because we did not observe any positive effect of this treatment. Incubation was done at 26°C.

For the dialysis system, 500 μl of reaction mixture contained 300 μl of the extract and the same ingredients as described above. The dialysis bag was immersed in 5 ml of a solution containing all described ingredients except for creatine kinase. The reaction was done at 23°C, and, every 24 h, 0.05 nmol of DHFR mRNA (or equivalent moles of the other mRNAs) and 50 μg of creatine kinase were supplemented. The dialysis solution was also replaced every 24 h. To confirm the longevity of the system, [^{14}C]leucine (the same concentration as above) was added into both reaction mixture and dialysis buffer at 52 h, then was incubated until 72 h (Fig. 4*C*). The autoradiogram of the gel was obtained by using a BAS-2000 phosphoimager (Fuji).

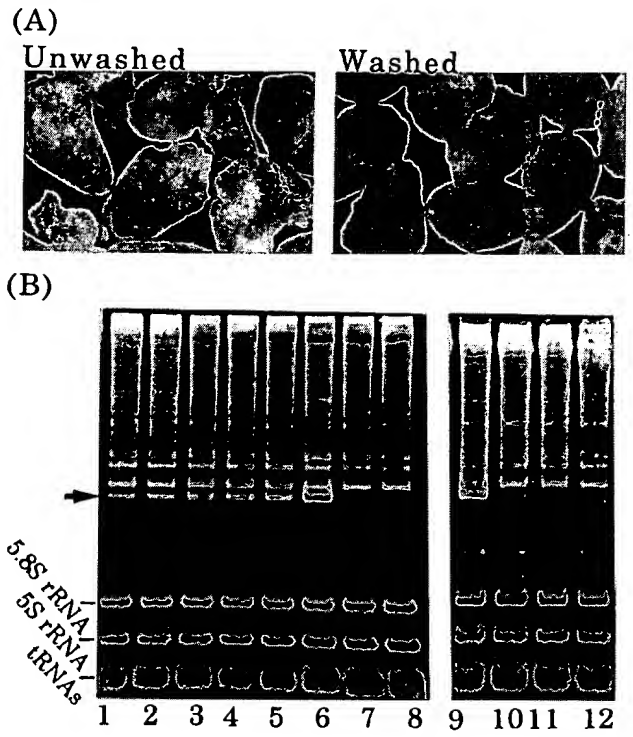

Fig. 1. Removal of tritin from embryos. Extracts were prepared from unwashed or washed embryos (*A*), and the depurination assay was performed (*B*). Translation mixtures prepared with the extract from unwashed embryos were incubated for 0, 1, 2, 3, and 4 h (*B*, lanes 1–5 respectively); mixtures with washed embryos were incubated for 0, 2, and 4 h (lanes 10–12, respectively). Isolated RNA was treated with acid/aniline, then was separated on 4.5% polyacrylamide gels. Additionally, RNA was directly extracted from embryos with guanidine isothiocyanate-phenol and was analyzed as above before (*B*, lane7) and after (*B*, lane 8) treatment with acid/aniline. For the fragment marker (*B*, lanes 6 and 9), incubation was carried out in the presence of gypsophilin (40), a highly active RIP from *Gypsophila elegance*; the arrow indicates the aniline-induced fragment.

Preparation of mRNA. Capped mRNA encoding DHFR was synthesized by *in vitro* transcription of linearized plasmid pSP65 carrying the gene under SP6 RNA polymerase promoter control (9). The transcript is 1,079 nucleotides long and consists of the sequence m^7GpppGAAUACACGGAAUUCGAGCUCG-CCCGGGAAAUCUCA*AUG* (the italicized sequence is the initiation codon) at its 5′ end, a 477-nt coding sequence, and a 3′ noncoding region of 565 nucleotides with a poly(A) tail of 100 adenosines (9). Coding sequences for GFP (717 nucleotides) (25) and luciferase (1,650 nucleotides) were cloned into the above plasmid in such a way that the 5′- and 3′-untranslated regions of DHFR were preserved. Capped TMV RNA (6,388 nucleotides) was transcribed from linearized plasmid pTLW3 carrying the genome under T7 RNA polymerase promoter control (26).

Analysis of Products and Their Enzymatic Activities. The amount of protein synthesized was determined as follows: Aliquots were withdrawn, and samples containing 1 μl of reaction mixture were separated on 12.5% SDS polyacrylamide gels (8% gels for TMV protein) or 12.5% native polyacrylamide gels (for GFP), then were stained with Coomassie brilliant blue. The product amount was estimated by densitometric scanning of the bands and comparison to standards. The standard samples were prepared by mixing a reaction mixture without mRNA with known amounts of standard proteins (DHFR, GFP, or luciferase) before loading onto the gel. Because pure, authentic, 126-kDa TMV protein is not available, the amount of this protein was estimated with less accuracy by calculating its relative amount compared with molecular markers included as internal standards by using average 105- and 160-kDa band intensities. The amount of DHFR was confirmed by determining the amount of methotrexate-agarose column purified protein, and its activity was measured colorimetrically as described (9). Luciferase activity was determined by using a commercial kit and a liquid scintillation counter as described (28). The specific activities of recombinant luciferase and the synthesized protein were 3.4×10^5 and 5.1×10^6 cpm/pg, respectively. Semiquantitative measurement (28) of GFP activity on the native gel was carried out by using a UV-illuminator (Dark Reader, Clare Chemical Research, Denver) with a wavelength of 400–500 nm. Subsequent scanning of photographs of the UV images and comparison of the intensities of the bands to those of the recombinant protein revealed that the translation product had more activity than the standard by a factor of 1.4.

Results and Discussion

Removal of Contaminants such as Tritin from Wheat Embryos Leads to a More Active Cell-Free Protein Synthesis System. Since the first report of solvent flotation for the enrichment of viable, intact embryos from wheat seeds by Johnston and Stern (29), this method has commonly been used for the preparation of wheat embryos. We first addressed the possibility of a tritin contamination originating from endosperm as the reason for the instability of wheat germ cell-free systems. If wheat germs are isolated from dry wheat seeds by conventional procedures (27), microscopic examination reveals that the sample contains embryos as well as some white material and a number of white and brownish granules (Fig. 1*A*). Analysis of ribosomal RNAs from a protein synthesis reaction prepared from such a sample showed that depurination of ribosomes occurs, contradicting earlier reports (20–22) (Fig. 1*B*). After 4 h of incubation, 24% of the ribosome population had been depurinated, as judged by the aniline-dependent formation of a specific RNA fragment (Fig. 1*B*,

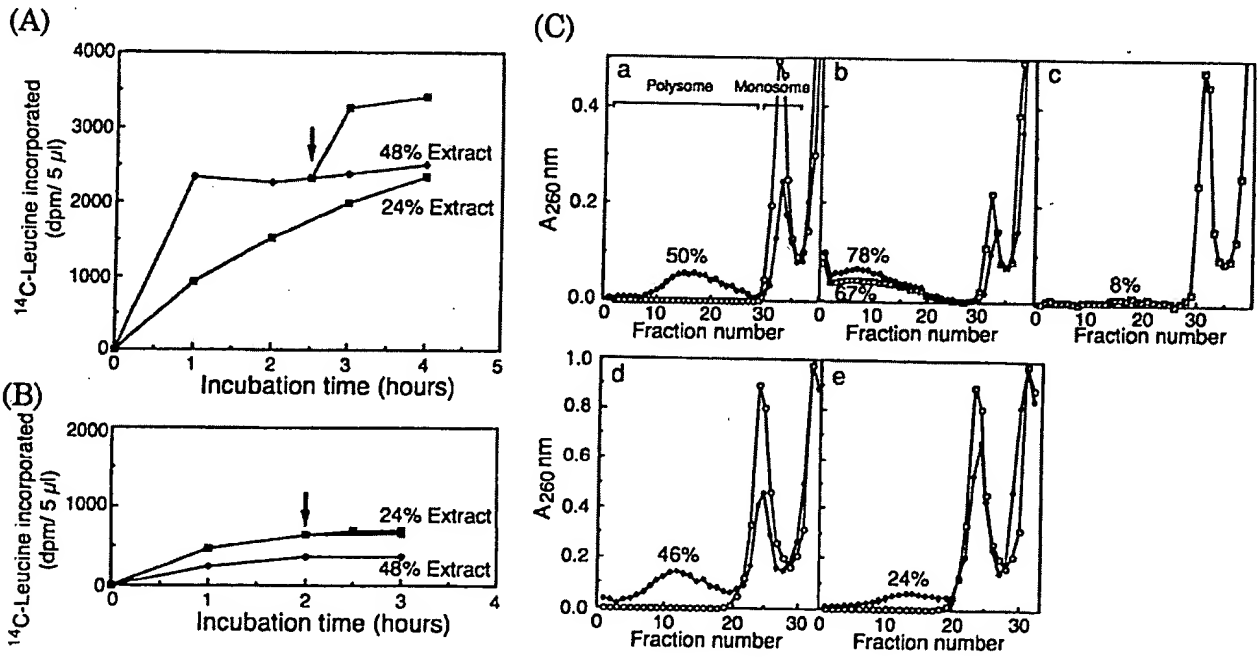

Fig. 2. Protein synthesis with an extract prepared from washed embryos. The batch system contains either 12 μl (24%) or 24 μl (48%) of extracts from washed (*A*) or unwashed (*B*) wheat embryos. Protein synthesis was measured as hot trichloroacetic acid insoluble radioactivity. Arrows show addition of substrates. *C* shows the polysome profiles of 15 μl of reaction mixture aliquots loaded onto a linear 10% to 45% sucrose gradient in 25 mM Tris·HCl (pH 7.6), 100 mM KCl, and 5 mM $MgCl_2$. After centrifugation, fractions were collected from the bottom of the tubes and were measured at 260 nm as described (24). Incubation times were 0 h (open circles in *a*), 1 h (closed circles in *a*), and 2 h (*b*) in the absence (open circles in b) or presence (closed circles in *b*) of 0.4 μM cycloheximide. In *c*, the translation system prepared from unwashed embryos was incubated for 2 h. In *d* and e, aliquots from the dialysis system were withdrawn after 48 and 60 h and were incubated in the presence of 0.4 μM cycloheximide for another 60 min at 26°C (closed circles). Similar analyses of the samples were carried out in the absence of mRNA (*d* and e, open circles) as negative controls.

arrow). Furthermore, even at the start of the incubation, 7% of the population had already been depurinated. The site of depurination was confirmed by direct sequencing of the fragment to be in the universally conserved sarcin/ricin domain of 28S rRNA (data not shown). When RNA was extracted directly from embryos by guanidine isothiocyanate-phenol, little formation of the aniline-induced fragment was observed (Fig. 1*B*, lanes 7 and 8). Thus, depurination must have occurred during the extract preparation and then continued during the protein synthesis reaction.

The observed extent of depurination constitutes a considerable damage to protein synthesis because inactivation of any one ribosome among the actively translating ribosomes on an mRNA results in blockage of the respective polyribosome and cessation of translation (16). Attempts were made to neutralize the depurinating enzyme with synthetic RNA aptamers that tightly bind to the RIP (30), but these attempts failed. Instead, careful selection and subsequent extensive washing of the embryos yielded better results. These embryos had few contaminants (Fig. 1*A Right*), and when the depurination assay was performed, no aniline-induced cleavage was detectable (Fig. 1*B*, lanes 10–12), indicating minimal, if any, depurination during preparation as well as incubation.

As shown in Fig. 2, the cell-free system prepared from washed embryos has much higher translational activity than the conventional system (compare Fig. 2 *A* and *B*). When programmed with mRNA coding for DHFR, it has almost linear kinetics in DHFR synthesis over 4 h in a system containing 24% extract, as opposed to the regular system, which ceased to function after 1.5 h. When the content of washed extract in the reaction volume was increased to 48%, amino acid incorporation occurred initially at a rate twice that with 24% extract, but then stopped after 1 h. However, this halting was caused by a shortage of substrates rather than an irreversible inactivation of ribosomes or factors necessary for translation: Addition of amino acids, ATP, and GTP after cessation of the reaction (Fig. 2 *A* and *B*, arrows) restarted translation with kinetics similar to the initial rate. In contrast, if conventional extract was added to 48%, protein synthesis actually decreased compared with the 24% extract reaction. Furthermore, the halting of protein synthesis in the

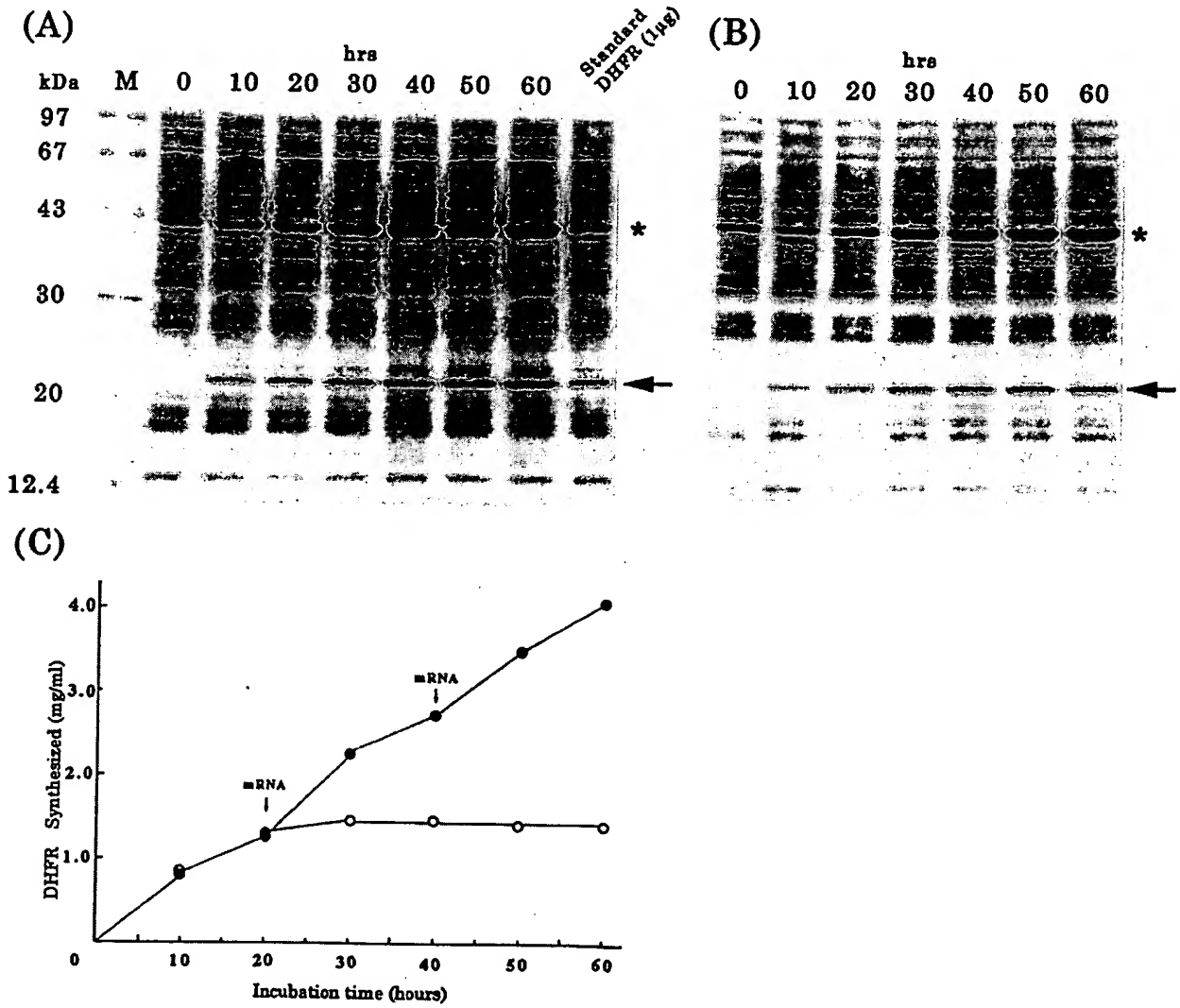

Fig. 3. Protein synthesis in the dialysis system. (*A* and *B*) Coomassie blue-stained SDS polyacrylamide gels showing DHFR synthesis with (*A*) or without (*B*) addition of new mRNA. Arrows and asterisks mark DHFR and creatine kinase, respectively. The standard sample was prepared by mixing a reaction mixture without mRNA with known amounts of DHFR before loading onto the gel. (*C*) Amounts of DHFR synthesized as determined from densitometric scans of the gels in *A* (closed circles) and *B* (open circles).

reaction with 24% extract could not be reversed by the addition of more substrate, indicating an irreversible damage by contaminants from endosperm (Fig. 2*B*).

High protein synthesis activity of the system with washed embryos can also be demonstrated by sucrose density gradient analysis (Fig. 2*C*). Significant formation of polysomes was observed after 1 h of incubation, and at 2 h a shift to heavier polysomes with a concomitant decrease of 80S monosomes was seen (Fig. 2*C a* and *b*). In the presence of low concentrations of cycloheximide polysome formation is a measure of translational initiation (31). A concentration of cycloheximide of 0.4 μM reduced the incorporation of [^{14}C]leucine to 21% of the control (data not shown) and resulted in an accumulation of large polysomes, with 78% of ribosomes in polysomes (open circle in Fig. 2*C b*). A similar analysis of cell-free reactions prepared with regular extracts (27), but done in the absence of cycloheximide, did not show significant polysome formation (Fig. 2*C c*). The high efficiency of our system, therefore, can be attributed to at least two factors: first, high initiation, elongation, and termination rates (efficient usage and recycling of ribosomes); and second, low endogenous ribonuclease activity (retention of heavy polysomes for prolonged time).

There is an additional explanation for the dramatic improvement of protein synthesis after washing of the embryos. Thionins are a group of small basic and cysteine-rich proteins, originally purified as antifungal proteins from a variety of plants, including wheat seeds (32). Wheat γ-thionin is known to be in the endosperm of seeds (33), and, recently, Brummer *et al.* have shown in a wheat germ translation system that α- and β-thionin from barley endosperm are potent inhibitors of protein synthesis initiation (34). In addition, several ribonucleases have been reported in the endosperm of the seeds (35). Thus, it is possible that the washing of the embryos resulted in elimination of thionin and ribonucleases as well as tritin.

The Continuous-Flow Cell-Free System on a Preparative Scale. After establishing a procedure for the preparation of highly active wheat embryo extract, we addressed its possible application for the large scale production of protein. For this purpose, we chose a dialysis system because of its continuous supply of substrates and continuous removal of small byproducts (12). With DHFR mRNA as template, protein synthesis worked efficiently, as demonstrated by a Coomassie blue stained gel (Fig. 3*A*, arrow). Densitometric quantitation as well as a direct determination of purified DHFR revealed that the reaction proceeded up to 60 h, yielding 4 mg of enzyme in a 1-ml reaction (Fig. 3*C*). This yield was achieved when the system was supplemented with fresh mRNA every 24 h; without the addition of fresh mRNA, the reaction ceased after 24 h and yielded 1 mg of DHFR (Fig. 3 *B* and *C*, open circles). When aliquots of the reaction mixtures were withdrawn after 48 and 60 h and then were incubated in the presence of a low dose of cycloheximide for an additional 1 h, sucrose gradient centrifugation revealed polysome formation (Fig. 2*C d* and *e*). This is a direct indication of a robust system with high translational activity. The product has a similar specific activity as the authentic enzyme, 15.3 vs. 19.1 units/mg (9).

As shown in Fig. 4, the system also synthesized proteins of higher molecular weight in a preparative scale: 1.1 mg of luciferase (65 kDa), 1.2 mg of GFP (45 kDa). These proteins had the same or even higher specific activity compared with commercially available recombinant forms (Fig. 4 *A* and *B*; see *Materials and Methods*). Furthermore, the 126-kDa replicase of TMV, a major genome product (36) during infection, was produced with a yield of as much as 0.6 mg (Fig. 4). The synthesis proceeded for up to 72 h, as shown by the increase in intensity of the Coomassie brilliant blue-stained bands. This point was confirmed by autoradiography and analysis of amino acid incorporation: [^{14}C]leucine was added after 52 h, samples were withdrawn at 60 and 72 h, and the samples were analyzed by SDS gel electrophoresis and autoradiography (Fig. 4). Densitometric quantitation of the bands showed linear synthesis: The photostimulated luminescence of the sample after 8 h of synthesis (at the 60-h time point) was 186, and after 20 h (at the 72-h point) it was 465, even though the rate of protein synthesis as measured by leucine incorporation was 21% of the rate at the beginning of incubation. This is another direct evidence of the robustness of

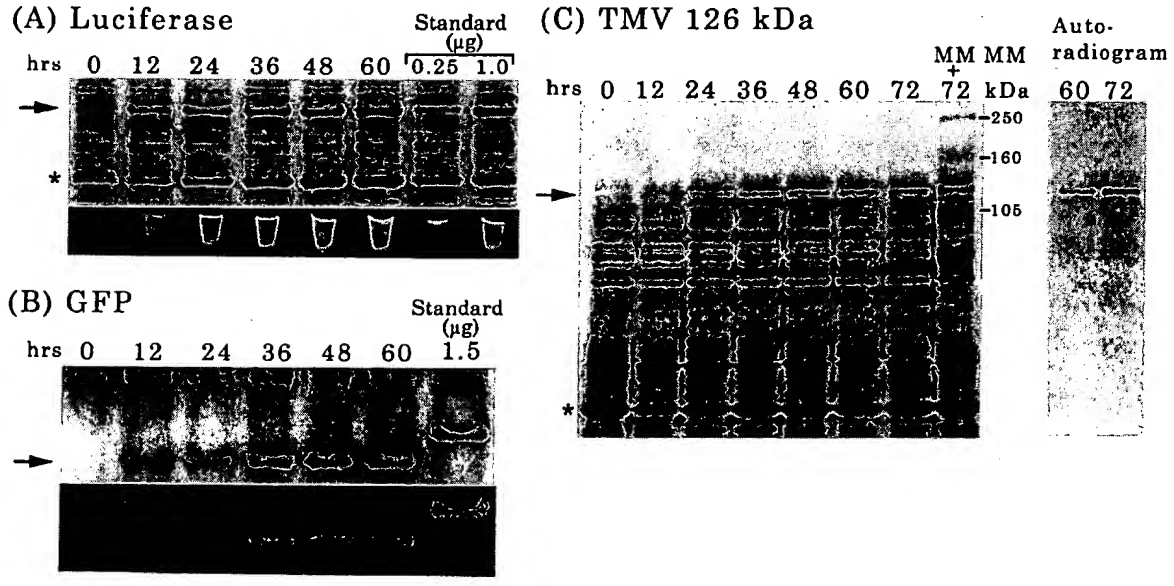

Fig. 4. Synthesis of luciferase (*A*), GFP (*B*), and 126-kDa TMV protein (*C*) in the dialysis system. Samples were analyzed as described in *Materials and Methods*. The standard samples were prepared by mixing a reaction mixture without mRNA with known amounts of luciferase or GFP before loading onto the gel. For the autoradiogram in *C*, [^{14}C]leucine was added at 52 h, and samples were withdrawn after an additional 8 h (60 h total) or 20 h (72 h). Authentic GFP migrates slower than the cell-free product on the native gel, which is attributable to different amino acid compositions because both proteins work as a monomer form. Products and supplemented creatine kinase are marked with arrows and asterisks, respectively.

the system and its efficiency in synthesizing even a 126-kDa protein for 3 days.

The structures of 5′- and 3′-untranslated regions are important for the efficiency of initiation and termination and also for the stability of mRNA (37). The mRNA constructs used here were not optimized in this respect, and we believe that the yields in our experiments do not, therefore, necessarily reflect maximum capacity. Efficient mRNA translation and its regulation requires a series of protein–mRNA and protein–protein interactions (37), and Wells *et al.* have recently shown the circularization of mRNA *in vitro* (38). Our method provides, in addition to its protein synthesis capacity, the opportunity to study translation itself, including the phenomenon of circular mRNA or the characterization of untranslated regions of mRNA in terms of efficient initiation or stability.

We show here that removal of endosperm contaminants, which contain protein synthesis inhibitor(s), from the embryo fraction improves protein synthesis in a wheat germ cell-free system. The improvement likely is caused by increased translational activity resulting from elimination of inhibitors of initiation (e. g. the thionins) and ribonucleases, as well as elimination of the RIP tritin. It is generally believed that cell-free translation systems are inherently unstable, but our results demonstrate the opposite: The translational apparatus appears to be very stable, *in vitro* and presumably also *in vivo*. We believe that our results shed light on the biological function of the nearly ubiquitous plant RIPs. We propose that plants acquired during evolution a suicide system useful to prevent larger damage and that because of its stability the translational machinery is the most important target of a suicide system. Viral attack would be one instance in which this suicide mechanism is employed. Ribosomes are a popular target of antibiotics also, emphasizing their central role in cell metabolism. The observed high stability of the translational apparatus might be an essential requirement for the evolution of life: Certain basic physiological processes such as protein synthesis might be required to function even in adverse conditions.

It is likely that the strategy that we followed to improve the wheat cell-free system, i.e., the inactivation of the translational suicide system, is successful with other systems as well. For instance, the widely used cell-free system from *E. coli* contains high ribonuclease activity and is hampered by a low efficiency in the translation of large mRNAs. Because of significant levels of template degradation, *E. coli* systems are limited when selecting large polypeptides for polysome display.

Our protein synthesis system has several advantages compared with existing systems in addition to its high efficiency: As a eukaryotic system, it is more amenable to the production of eukaryotic proteins from their natural mRNAs: i.e. no cDNA modification is needed; the system can produce high molecular weight proteins; because of little template degradation, it is useful for polysome display (39); and proteins that would normally interfere with cell physiology can be synthesized. Additionally, it should be a useful tool in the study of translation itself.

This work was supported by The Japan Society for the Promotion of Science Grant JSPS-RFTF 96100305 (to Y.E.).

1. Golf, S. A. & Goldberg, A. L. (1987) *J. Biol. Chem.* **262,** 4508–4515.
2. Chrunyk, B. A., Evans, J., Lillquist, J., Young, P. & Wetzel, R. (1993) *J. Biol Chem.* **268,** 18053–18061.
3. Henrich, B., Lubitz, W. & Plapp, R. (1982) *Mol. Gen. Genet.* **185,** 493–497.
4. Kurland, C. G. (1982) *Cell* **28,** 201–202.
5. Pavlov, M. Y. & Ehrenberg, M. (1996) *Arch. Biochem. Biophys.* **328,** 9–16.
6. Roberts, B. E. & Paterson, B. M. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 2330–2334.
7. Spirin, A. S., Baranov, V. I., Ryabova, L. A., Ovodov, S. Yu. & Alakhov, Yu. B. (1988) *Science* **242,** 1162–1164.
8. Baranov, V. I., Morozov, I. Yu., Ortlepp, S. A. & Spirin, A. S. (1989) *Gene* **84,** 463–466.
9. Endo, Y., Otsuzuki, S., Ito, K. & Miura, K. (1992) *J. Biotechnol.* **25,** 221–230.
10. Kigawa, T. & Yokoyama, S. (1991) *J. Biochem.* **110,** 166–168.
11. Kim, D.-M., Kigawa, T., Choi, C.-Y. & Yokoyama, S. (1996) *Eur. J. Biochem.* **239,** 881–886.
12. Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T. & Yokoyama, S. (1999) *FEBS Lett.* **442,** 15–19.
13. Kawarasaki, Y., Kawai, T., Nakano, H. & Yamane, T. (1995) *Anal. Biochem.* **226,** 320–324.
14. Endo, Y., Mitsui, K., Motizuki, M. & Tsurugi, K. (1987) *J. Biol. Chem.* **262,** 5908–5912.
15. Endo, Y. & Tsurugi, K. (1987) *J. Biol. Chem.* **262,** 8128–8130.
16. Wool, I. G., Glück, A. & Endo, Y. (1992) *Trends Biochem. Sci.* **17,** 266–269.
17. Barbieri, L., Battelli, M. G. & Stirpe, F. (1993) *Biochim. Biophys. Acta* **1154,** 237–282.
18. Taylor, S., Massiah, A., Lomonossoff, G., Robert, L. M., Lord, J. M. & Hartely M. R. (1994) *Plant J.* **5,** 827–853.
19. Ready, M. P., Brown, D. T. & Robertus, J. D. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 5053–5056.
20. Massiah, A. J. & Hartely, M. R. (1995) *Planta* **197,** 633–640.
21. Stewart, T. S., Hruby, D. E., Sharma, O. K. & Robert, W. K. (1977) *Biochim. Bipphys. Acta* **479,** 31–38.
22. Taylor, B. E. & Irvin, J. D. (1990) *FEBS Lett.* **273,** 144–146.
23. Yoshinari, S., Koresawa, S., Yokota, S., Sawamoto, H., Tamura, M. & Endo, Y. (1997) *Biosci. Biotechnol. Biochem.* **61,** 324–331.
24. Hase, M., Endo, Y. & Natori, Y. (1982) *J. Biochem.* **91,** 1457–1465.
25. Chiu, W.-L., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H. & Sheet, J. (1996) *Curr. Biol.* **6,** 325–330.
26. Hamamoto, H., Sugiyama, K., Nakagawa, N., Hashida, E., Matsunaga, Y., Takemoto, S., Watanabe, Y. & Okada, Y. (1999) *Bio/Technology* **11,** 930–932.
27. Erickson, A. H. & Blobel G. (1983) *Methods Enzymol.* **96,** 38–50.
28. Aleen, R. C. (1986) *Methods Enzymol.* **133,** 449–493.
29. Johnston, F. B. & Stern, H. (1957) *Nature (London)* **179,** 160–161.
30. Hirao, I., Yoshinari, S., Yokoyama, S., Endo, Y. & Ellington A. D. (1997) *Nucleic Acids Symp. Ser.* **37,** 283–284.
31. Lodish, H. F., Housman, D. & Jacobsen, M. (1971) *Biochemistry* **10,** 2348–2356.
32. Bohlmann, H. (1994) *Crit. Rev. Plant Sci.* **13,** 1–16.
33. Colilla, F. J., Rocher, A. & Mendez, E. (1990) *FEBS Lett.* **270,** 191–194.
34. Brummer, J., Thole, H. & Kloppstech, K. (1994) *Eur. J. Biochem.* **219,** 425–433.
35. Matsushita, S. (1959) *Memoirs Res. Inst. Food Sci., Kyoto Univ.* **19,** 1–4.
36. Dawson, W. O. (1992) *Virology* **186,** 359–367.
37. Sachs, A. B., Sarnow, P. & Hentze, M. W. (1997) *Cell* **89,** 831–835.
38. Wells, S. E., Hillner, P. E., Vale, R. D. & Sachs, A. B. (1998) *Mol. Cell* **2,** 135–140.
39. Mattheakis, L. C., Bhatt, R. R. & Dower, W. J. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 9022–9026.
40. Yoshinari, S., Koresawa, S., Yokota, S., Sawamoto, H., Tamura, M. & Endo, Y. (1997) *Biosci. Biotechnol. Biochem.* **61,** 324–331.

The EMBO Journal vol.9 no.5 pp.1665 - 1672, 1990

Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure

Wayne A.Hendrickson[1,2], John R.Horton[2] and David M.LeMaster[3,4]

[1]Howard Hughes Medical Institute, [2]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, and [3]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511, USA

[4]Present address: Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL 60208, USA

Communicated by G.N.Cohen

An expression system has been established for the incorporation of selenomethionine into recombinant proteins produced from plasmids in *Escherichia coli*. Replacement of methionine by selenomethionine is demonstrated at the level of 100% for both T4 and *E.coli* thioredoxins. The natural recombinant proteins and the selenomethionyl variants of both thioredoxins crystallize isomorphously. Anomalous scattering factors were deduced from synchrotron X-ray absorption measurements of crystals of the selenomethionyl proteins. Taken with reference to experience in the structural analysis of selenobiotinyl streptavidin by the method of multiwavelength anomalous diffraction (MAD), these data indicate that recombinant selenomethionyl proteins analyzed by MAD phasing offer a rather general means for the elucidation of atomic structures.

Key words: crystallography/protein structure/selenium/selenomethionine/thioredoxin

Introduction

The decisive molecular images obtained from crystal structures often provide powerful insight into biological activity. These images have traditionally been hard won, however, as macromolecular crystallography is generally arduous and time consuming. Recent advances in instrumentation and methodology have accelerated the pace substantially; but, nevertheless, two steps in the process remain especially problematic: crystallization and the phase problem. There is still more art than science in growing suitable crystals, although prospects are improved by the availability of clean protein stocks in abundance from recombinant expression systems. Molecular biology techniques can also aid in evaluating phase angles for the diffracted X-ray waves, and we report here on a new approach for introducing suitable diffraction labels into proteins.

One must know both the amplitude and phase for each of thousands of diffracted waves from a macromolecular crystal in order to reconstruct an image, but only amplitudes can be recovered from standard diffraction measurements. This presents the central conceptual obstacle in crystallography—the phase problem. Phase evaluation for new macromolecular structures traditionally has been based on the analysis of isomorphous replacements with heavy atoms, although other sources of phasing such as molecular averaging or anomalous scattering are often used in supplement. This, the method of multiple isomorphous replacement (MIR), entails the trial-and-error preparation of heavy-atom derivatives and it is frequently limited by lack of isomorphism.

Recently, an alternative approach has been devised for *de novo* phasing of macromolecular crystal structures. This is the method of multiwavelength anomalous diffraction (MAD) (Hendrickson, 1985) which exploits the scattering effects of resonance between X-rays and bound atomic orbitals as indicated in Figure 1. Qualitatively, MAD experiments can be thought of as *in situ* isomorphous replacements generated by the variation in scattering strength that accompanies change of wavelength. The MAD method does require the special properties of synchrotron radiation, but it has advantages in that isomorphism is perfect, all data can be measured from a single crystal, and the analysis is algebraically exact (Karle, 1980). Several recent applications demonstrate the effectiveness of MAD phasing (Guss *et al.*, 1988; Hendrickson *et al.*, 1988; Krishna Murthy *et al.*, 1988; Hendrickson *et al.*, 1989).

The MAD method requires the presence of a few heavier elements as resonance centers. The anomalous scattering from the lighter atoms (H, C, N and O) in proteins is inappreciable at typically achievable wavelengths (0.5–3.0 Å), and although sulfur anomalous scattering has been used for protein phasing in special cases (Hendrickson and Teeter, 1981) its resonance energy is ordinarily inaccessible for MAD experiments. Either intrinsic metal centers (e.g. Fe or Zn) or introduced heavy atoms (e.g. Hg, Au or Gd) are suitable and they can produce dramatic effects. However, metalloproteins or chemical derivatives are not always at hand. Systematic biological incorporation of selenomethionine in place of methionine residues in proteins offers the possibility of generality. Cowie and Cohen (1957) showed that a strain of *E.coli* made auxotrophic for methionine could grow for 100 generations in selenomethionine, suggesting that all proteins needed to sustain life in this organism can function well when fully substituted with selenomethionine. Our experiments in the MAD phasing of selenolanthionine (Hendrickson, 1985) and of selenobiotinyl streptavidin (Hendrickson *et al.*, 1989) demonstrate that selenium is very effective as a center for MAD phasing.

In this study, we have developed a system for expressing recombinant selenomethionyl proteins in *E.coli*. Selenomethionyl thioredoxins produced from this system have been crystallized and characterized. A theoretical evaluation based on anomalous scattering factors measured from selenomethionyl *E.coli* thioredoxin indicates that MAD analysis of selenomethionyl proteins should provide a broadly applicable solution to the phase problem.

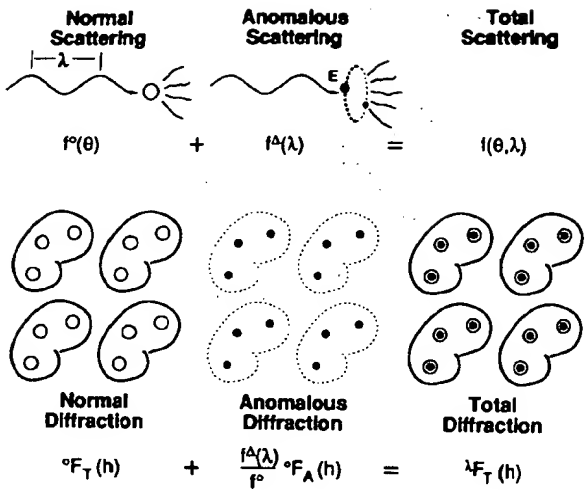

Fig. 1. Schematic description of anomalous scattering from atoms in macromolecular crystal structures. The upper portion illustrates the normal and anomalous components of atomic scattering. Normal scattering of X-rays by matter arises from radiation emitted by electrons that have been excited into free vibration by an incident X-ray wave. The amplitude of normal atomic scattering, $f^\circ(\theta)$ depends on scattering angle, θ, but it is independent of wavelength, λ. However, when the energy, E, associated with X-rays of wavelength λ approaches an energy level of electrons bound in atomic orbitals, the induced vibrations resonate with the intrinsic atomic oscillations. This affects the amplitude and the phase of scattering. The incremental anomalous scattering factor, $f^\Delta(\lambda)$, depends very strongly on wavelength but is essentially independent of scattering angle. The lower portion illustrates the components of diffraction from a crystalline array of molecules. Each molecule in a protein crystal might contain a few thousand atoms (such as C, N and O) that only scatter X-rays normally and few additional atoms (such as Se) that produce anomalous (resonant) scattering. The normal component of anomalous centers is symbolized by open circles and the anomalous component is represented by solid circles. The preponderant light atoms are not represented individually here. The total diffraction from such a crystal is described by the structure factor, $^\lambda F_T(h)$, for each reflection h as measured at a particular wavelength λ. The normal scattering component of this diffraction, which has components from all atoms in the structure, is given by $^\circ F_T(h)$. Knowledge of this wavelength independent component of diffraction, $^\circ F_T = |^\circ F_T| \exp(i^\circ\phi_T)$, including its phase, suffices by Fourier transformation to produce an image of the crystal structure. Only the heavier atoms, usually few in number, contribute to anomalous diffraction, and this component contains all wavelength dependence. In the case of a single kind of anomalous scatterer (as in selenomethionyl proteins) this component is related to the normal scattering contributions from these atoms, $^\circ F_A(h)$, by the ratio of anomalous to normal scattering factors, $f^\Delta(\lambda)/f^\circ(\theta)$.

Through an algebraic analysis of total diffraction data measured accurately at a few wavelengths (usually 3–5), it is possible to extract the essential components $|^\circ F_T|$, $|^\circ F_A|$ and $\Delta\phi = {}^\circ\phi_T - {}^\circ\phi_A$ from the set of $|^\lambda F_T|$ measurements (Karle, 1980; Hendrickson, 1985). This is done for each of the several thousand reflections. Then the $|^\circ F_A|$ values are used to deduce the structure of anomalous scattering centers, and from this structure $^\circ\phi_A$ values are computed. These together with the fitted $\Delta\phi$ values serve to specify $^\circ\phi_T$ and thereby solve the phase problem. The distributions of electron density are synthesized by Fourier transformation of $|^\circ F_T| \exp(i\ ^\circ\phi_T)$.

Results

Selenomethionine tolerant met⁻ strain

Different met⁻ strains of *E.coli* tolerate the substitution of selenomethionine in varying degrees. After testing the growth characteristics of a few methionine auxotrophs in selenomethionine medium, one of us (D.M.LeMaster) constructed a new auxotrophic strain, DL41, by introducing a lesion in the *metA* gene. This strain grows as vigorously as its parent MG1655 (Guyer *et al.*, 1981), a viable F^-, λ^- strain derived from *E.coli* K12. It can double in 80–85 min at 33°C on a defined glucose medium (LeMaster and Richards, 1985) supplemented by selenomethionine in place of methionine (as compared with 50–55 min in the methionine medium), and it will grow to an OD_{600} of >4.0 in a shaker bath. Cell growth in selenomethionine is somewhat variable, however. For example, the doubling time in shaker flasks at 31°C was found to be 160 min.

Recombinant selenomethionyl thioredoxin expression

Our initial tests in the production and crystallographic analysis of selenomethionyl proteins have been carried out on thioredoxins since a suitable expression system was already available. This system simultaneously produces both *E.coli* and bacteriophage T4 thioredoxins (LeMaster and Richards, 1988) and these molecules have concentrations of methionine suitable for MAD phasing: one residue in 108 for the *E.coli* protein (Holmgren, 1968) and three in 87 for the T4 protein (Sjöberg and Holmgren, 1972). The expression plasmid for this system, pDL59, is under control of the P_L promoter from bacteriophage λ, and it was constructed to be genetically portable for use in isotopic labeling of specific amino acids for NMR studies. Thus, the temperature-sensitive C_{I857} λ repressor gene and the antiterminator N gene were cloned into the plasmid along with the *trxA* gene for *E.coli* thioredoxin (Lunn *et al.*, 1984) and the *nrdC* gene for T4 phage thioredoxin (LeMaster, 1986). The pDL59 plasmid was introduced into the DL41 bacterial strain and this system was used to produce both selenomethionyl thioredoxins as well as the natural recombinant proteins. Starter cultures for selenomethionyl protein production were grown in methionine containing medium since stationary phase bacteria do not tolerate selenomethionine very well. The same purification procedures as described for the natural proteins (LeMaster and Richards, 1985) were used for purifying the selenomethionyl proteins. A yield of 65 mg of *E.coli* selenomethionyl thioredoxin and 20 mg of T4 selenomethionyl thioredoxin was obtained from 2.5 l of fermenter culture.

The level of selenomethionine substitution in the purified proteins was estimated by amino-acid composition analysis. Selenomethionine decomposes under the acid hydrolysis conditions used in the analysis (Shepherd and Huber, 1969), and thus it is the absence of methionine that is monitored. Comparable portions of the chromatograms for natural and selenomethionyl T4 thioredoxins are shown in Figure 2. Analyses made with normal loading (~10 μg) gave a yield of 2.5 and 1.0 residues of methionine per molecule from the natural T4 and *E.coli* proteins, respectively, whereas reproducibly there was no detectable methionine in the selenomethionyl proteins. When analyzed under overloaded conditions (30–40 μg), the T4 and *E.coli* selenomethionyl thioredoxins gave 0.04 and 0.02 residues of methionine per molecule (i.e. 1.3% and 2.0% of the methionine sites). These levels are well within the root mean square (r.m.s.) deviation of 0.09 residues between duplicates and a 0.61 residue r.m.s. discrepancy between analytical and theoretical compositions. Thus, within the limits of the sensitivity of our analysis the replacement of methionine by selenomethionine is complete in these experiments. Positive evidence for the selenomethionine substitution is provided by X-ray absorption spectra described below.

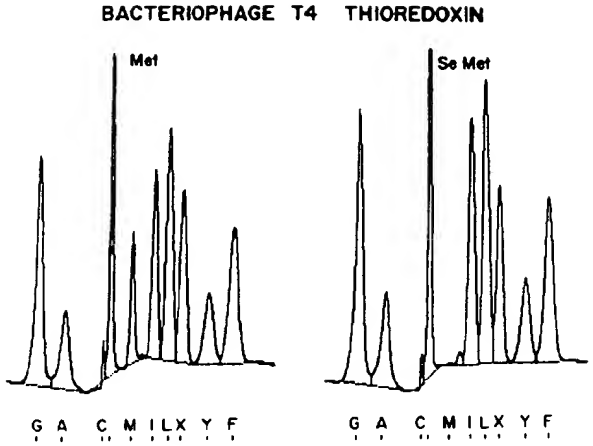

Fig. 2. Portions of the amino acid composition chromatograms for natural and selenomethionyl T4 thioredoxin. The respective elution profiles of absorbances of ninhydrin derivatives of the amino acids are shown with positions marked by the one-letter code for amino acids. Valine is in the unmarked peak following cysteine and X indicates the location of the norleucine standard.

Crystallization of selenomethionyl thioredoxins

Crystal structure analyses have been reported previously both for *E.coli* thioredoxin (Holmgren *et al.*, 1975) and for T4 thioredoxin (Söderberg *et al.*, 1978), and the structure of the *E.coli* protein has recently been refined at high resolution (Katti *et al.*, 1990). In this study we have attempted to crystallize the selenomethionyl proteins under conditions similar to those used with the natural methionyl thioredoxins. Pertinent characteristics of resulting crystals are summarized in Table I. T4 thioredoxin was crystallized at 4°C from ethanol in the presence of cadmium, but by vapor diffusion in hanging drops rather than in microdiffusion capillaries as before (Sjöberg and Söderberg, 1976). Nevertheless, the natural and selenomethionyl crystals are nearly isomorphous. Crystals of *E.coli* thioredoxin were grown in the presence of copper, but using polyethylene glycol (PEG) rather than 2-methyl-2,4-pentanediol (MPD) as the precipitating agent and at 20°C rather than 4°C. The PEG crystals differ markedly from the MPD crystals although they are clearly related. The natural recombinant protein and its selenomethionyl counterpart crystallize isomorphously in vapor diffusion against the same PEG concentration. This is evident from crystal morphology and X-ray diffraction patterns (Figure 3). The most noticeable difference in the crystallization of the two *E.coli* proteins is that crystals appear sooner in the selenomethionine case. This may reflect the somewhat reduced solubility of the variant protein—under the initial hanging-drop conditions (4% PEG 3350 and 1 mM Cu^{2+} at pH 4.2), a slight precipitate can be observed with selenomethionyl *E.coli* thioredoxin at 3.9 mg/ml, whereas under the same conditions but with the natural *E.coli* thioredoxin at 6.3 mg/ml the initial solution is clear.

Anomalous scattering factors

Since the efficiency of the MAD method derives from the variation of scattering strength with X-ray wavelength, it is essential that the anomalous scattering factors be known. These can be obtained from X-ray absorption spectra measured near the resonant orbital energy with synchrotron radiation and analyzed in relation to theoretical values (Cromer, 1983) that are computed for the elemental state. X-ray absorption spectra were measured from crystals of both *E.coli* and T4 thioredoxin, and these data were analyzed as before (Hendrickson *et al.*, 1988) to produce scattering factors. The resulting spectra are similar to those from other organic selenides including DL selenomethionine (unpublished results), selenolanthionine (Templeton and Templeton, 1988) and selenobiotinyl streptavidin (Hendrickson *et al.*, 1989). In particular, the selenium anomalous scattering is anisotropic which demonstrates that the selenomethionine is incorporated into thioredoxin in an ordered manner.

Anomalous scattering factors measured from one orientation of a selenomethionyl *E.coli* thioredoxin crystal are shown in Figure 4. These data were obtained with the high energy resolution of a Si111 double crystal monochromator and they exhibit a pronounced 'white line' of edge absorption. The strength of these sharp edge features enhances the phasing potential of the MAD method. Table II gives the locations and values of extrema in three nearly orthogonal orientations with respect to the polarized X-ray beam. A crystal of selenomethionyl T4 thioredoxin examined with radiation from a Ge111 monochromator gave a spectrum similar to that shown in Figure 4, but with somewhat lower resolution.

Theoretical feasibility for MAD phasing of selenomethionyl proteins

The ultimate test of the idea of direct determination of protein crystal structures by MAD phasing from selenomethionine is the actual analysis. We have experiments in progress on *E.coli* thioredoxin and other proteins. However, it is already clear by analogy with selenobiotinyl streptavidin that the MAD phasing of selenomethionyl proteins will be feasible provided that the methionine residues are ordered and appropriately abundant. The average frequency of occurrence of methionine residues in proteins is 1 in 59 residues (Dayhoff, 1978) and this compares favorably with the single selenium site per 126 residues in selenobiotinyl streptavidin.

Our analysis of anomalous diffraction data involves a least-squares optimization (Hendrickson, 1985) based on an exact algebraic analysis (Karle, 1980). An analysis of the prospective accuracy in a particular application of this procedure is difficult; however, the information content resides in diffracted intensity differences for which expected values can be estimated. Two such differences are pertinent and give mutually orthogonal phase information. These are the Bijvoet difference, $\Delta F_{\pm h} = |^{\lambda}F(\mathbf{h})| - |^{\lambda}F(-\mathbf{h})|$, and the dispersive difference, $\Delta F_{\Delta\lambda} = |^{\lambda_1}F| - |^{\lambda_2}F|$ where the structure factors, F, are as defined in the legend for Figure 1. Given the number of protein atoms, the number of anomalous centers, and the anomalous scattering factors it is possible to determine expected values for the respective diffraction ratios and thereby to estimate the relative phasing strength of a candidate for MAD phasing (Hendrickson *et al.*, 1985). Such estimates are presented in Table III for a number of pertinent actual and hypothetical situations. It is obvious that MAD phasing of prospective selenomethionyl proteins should be readily feasible since diffraction ratios are in excess of those that have proved adequate for related problems.

Table I. Characteristics of thioredoxin crystals

Molecule	Precipitant	Space group	Unit cell parameters				References
			a	b	c	β	
Natural *E.coli* thioredoxin	MPD, Cu^{2+}	C2	89.7 Å	51.1 Å	60.3 Å	113.5°	Holmgren *et al.* (1975)
Natural *E.coli* thioredoxin	PEG, Cu^{2+}	I2[a]	88.8	49.2	116.0	100.2	This work
Se met *E.coli* thioredoxin	PEG, Cu^{2+}	I2	89.2	49.6	115.6	100.2	This work
Natural T4 thioredoxin	Ethanol, Cd^{2+}	$P2_1$	54.1	45.9	40.8	99.4	Sjöberg and Söderberg (1976)
Se met T4 thioredoxin	Ethanol, Cd^{2+}	$P2_1$	53.9	45.8	39.0	100.4	This work

[a]The crystals are described here in the unconventional space group of I2 so that relatedness to the previously reported MPD crystals is evident. This unit cell is doubled in the c-direction but the corresponding additional reflections are relatively weak and are missing at low angles (d spacings >8 Å). The corresponding conventional cell in space group C2 has unit cell dimensions of a = 133.1 Å, b = 49.2 Å, c = 88.8 Å, β = 120.9°.

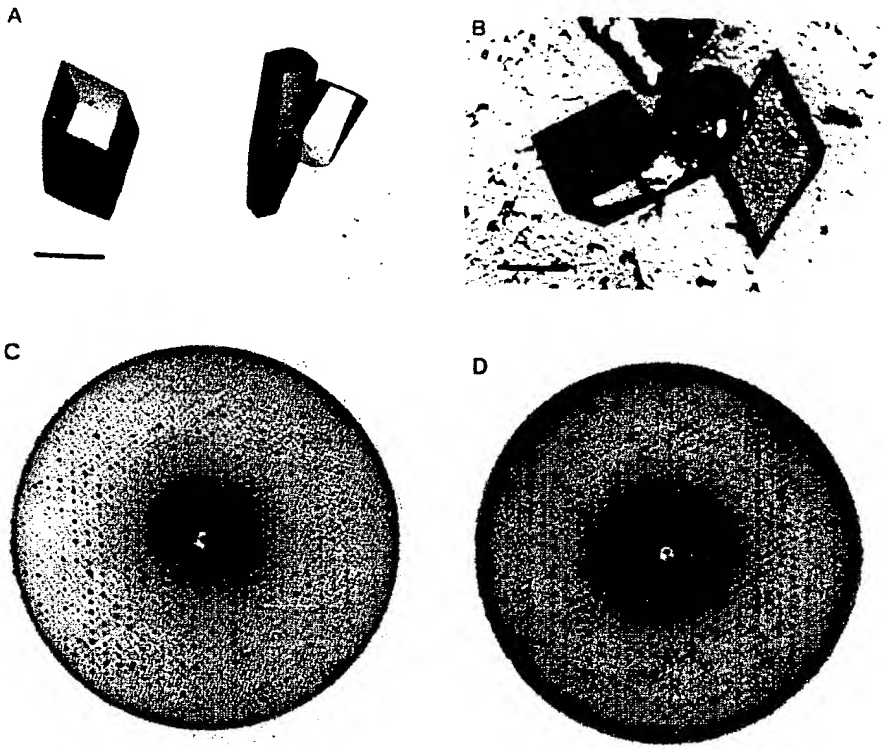

Fig. 3. Photographs of crystals and diffraction patterns of natural and selenomethionyl *E.coli* thioredoxin. (A) Type II crystals of the natural recombinant protein grown from PEG in the presence of Cu. The bar corresponds to 0.2 mm in length. (B) Type II crystals of the selenomethionyl protein. Protein used for this crystallization contained 10% methionine, but the crystals were essentially identical with those reported in Table I. The rounded object is a grain of sand used to induce nucleation. The bar corresponds to 0.2 mm. (C) A 15° precession photograph of the h0l zone of the diffraction pattern from a Type II crystal of natural thioredoxin. (D) A 15° precession photograph of the h0l pattern from a Type II crystal of selenomethionyl thioredoxin.

Discussion

Selenium chemistry and metabolism

The chemistry of selenium is similar to that of other elements in the oxygen series with which it shares the same outer electronic configuration. In this it has properties noticeably closer to sulfur than to its other neighbor tellurium. Indeed, selenium sufficiently resembles sulfur that it substitutes in a wide variety of organic compounds. Organic selenium chemistry is however distinctive (Odom, 1983). In particular, (i) single C–Se bonds are typically 0.14 Å longer than C–S bonds, (ii) organic selenium compounds are generally more reactive than their sulfur counterparts, owing presumably to lower bond strength, (iii) nucleophilic attack occurs more readily at selenium than at sulfur atoms, and (iv) selenium in comparison with sulfur is easier to oxidize from valence state (II) to (IV) but it is more difficult to oxidize from (IV) to (VI). Thus, selenides (R-Se-R′) and selenols (R-SeH) go more readily to selenoxides (R-SeO-R′) and seleninic acids (R-Se(O)OH) but less well on to selenones (R-SeO_2-R′) and selenonic acids (R-Se(O_2)OH) than happens for the sulfur analogs. In general, compounds with terminal selenium atoms (such as in selenols or selenoketones) degrade in air; whereas compounds with selenium bonded to two atoms (as in selenides such as selenomethionine) are relatively stable.

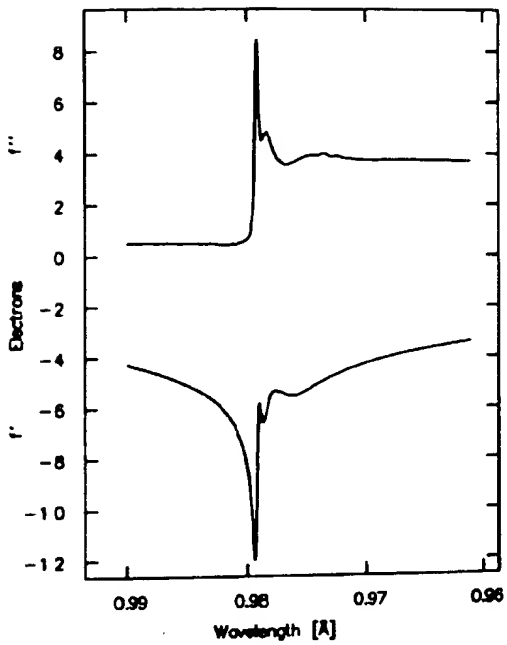

Fig. 4. Anomalous scattering factors near the absorption edge of selenium measured from a crystal of selenomethionyl *E.coli* thioredoxin. Here the electric vector E of the synchrotron X-ray beam was parallel with the b axis of the crystal.

Table II. Extrema of anomalous scattering from *E.coli* selenomethionyl thioredoxin

Feature	E || a*	E || b	E || c*	Average
Energy of f'_{min}	12 660 eV	12 659 eV	12 660 eV	12 659 eV
Minimal f'	−11.08	−11.83	−10.99	−11.20
Energy of f''_{max}	12 663 eV	12 662 eV	12 663 eV	12 662 eV
Maximal f''	7.82	8.18	7.93	7.88

The direction of X-ray polarization in the synchrotron beam is designated as E and direction in the crystal are along crystallographic axes in the real (b) and reciprocal lattice (a* and c*). The correspondence between energy levels and X-ray wavelength is $\lambda = 12.398/E$ (keV). Thus, 12 659 eV corresponds to 0.9794 Å and 12 662 eV corresponds to 0.9792 Å.

Selenium is both toxic and an essential element for most animal and bacterial life. The requirement for selenium is by virtue of enzymes that exploit selenocysteine as a strong nucleophile and through the occurrence of selenium in certain bacterial tRNA species (Stadtman, 1987). Its toxicity results from the facile substitution of selenium in sulfur metabolism with untoward effects due to the exceptional reactivity of some products. Sulfur metabolism is quite different in bacteria than it is in higher organisms. In bacteria, sulfur compounds derive from inorganic sulfur, typically sulfate, which is first used to produce cysteine and then methionine and other sulfur metabolites. In mammals, methionine ingested in foodstuffs is the sulfur source and cysteine derives from methionine. Yeast and fungi have pathways operating in both directions.

Selenomethionine is remarkably benign as a substitute for methionine. Cowie and Cohen (1957) interpreted their experiments, based on the strict absence of alternative sources, as demonstrating that 100% of the methionine in *E.coli* could be replaced by selenomethionine. However, direct physical assays were not performed, and assays for selenomethionyl azurin produced by *Pseudomonas aeruginosa* showed incorporation to be at most 94% complete (Frank *et al.*, 1985). Here we show specifically that in the case of recombinant thioredoxins the methionine residues can indeed be fully replaced. A number of early *in vitro* studies make this a plausible and expected result. First, the aminoacylation of methionyl tRNA by the methionyl-tRNA synthetases from *Sarcina lutea* (Hahn and Brown, 1967), *E.coli* and rat liver (McConnell and Hoffman, 1972) is indistinguishable when selenomethionine replaces methionine as a substrate. Secondly, and most stringently, selenomethionine is as good or better as a substrate for activation by S-adenosylmethionine synthetases from yeast and rat liver (Mudd and Cantoni, 1957), and the selenium analog is as effective as a methyl donor as is S-adenosylmethionine itself in choline biosynthesis by rat liver microsomes (Bremer and Natori, 1960). Finally, the activity and properties of selenomethionyl and natural proteins are found to be very similar, if not identical, as was first shown for β-galactosidase (Huber and Criddle, 1967).

Although selenomethionine is a faithful mimic of methionine, the same cannot be said for selenium metabolism generally. The growth of *E.coli* with selenite replacing sulfite as a sulfur source arrests after depletion of sulfur reserves (Cowie and Cohen, 1957). Moreover, the β-galactosidase produced by *E.coli* grown on a selenate medium proved to have selenomethionine incorporated at only the 53% level (Huber and Criddle, 1967). Interestingly, this purified protein contained no selenocysteine despite the necessary existence of selenocysteine as an intermediate in selenomethionine synthesis. Probably, selenocysteine (or an oxidized product) is not recognized by the cysteinyl tRNA synthetase.

Attributes of selenomethionine as a general phase vehicle

Until now the nearest thing to a general phasing vehicle for protein crystallography has been in the use of cysteinyl residues as targets for the preparation of isomorphous mercurial derivatives. In contemplating new opportunities for generality with the introduction of MAD methodology, several qualities for an ideal phasing label were considered. These considerations include both biochemical properties such as convenience of labeling production, degree of isomorphism with the parent molecule, and durability of the derivative product, and also diffraction concerns such as the spectral position of the anomalous element, prospective phasing power, and the relative abundance and mobility of target sites. The properties of selenomethionyl proteins relate quite favorably to these considerations.

A most significant attribute to selenomethionine is that biological incorporation into proteins obviates the complexities of chemical labeling. Equally important, the resulting selenomethionyl proteins appear to be essentially isostructural with their natural methionyl counterparts. That this is so could be anticipated from the isomorphism between crystal structures of D,L methionine and D,L selenomethionine (Rajeswaran and Parthasarathy, 1984) and the remarkable tolerance of protein structure to site-directed point mutation. The isomorphism between natural and selenomethionyl crystal structures that is demonstrated here for T4 and *E.coli* thioredoxins has now been extended to interleukin 1α (IL-1α)

Table III. MAD phasing strength from anomalous diffraction ratios

Molecule	Residues	N_A	N_P	Maximal Bijvoet ratio	Maximal dispersive ratio	Source of scattering factors
E.coli thioredoxin	108	1 Se	830	5.8%	3.5%	This work
T4 thioredoxin	87	3	670	11.1%	6.7%	This work
Hypothetical[a]	300	1	2310	3.5%	2.1%	This work
Hypothetical[a]	300	5	2310	7.7%	4.7%	This work
Hypothetical[a]	300	10	2310	10.9%	6.6%	This work
Hypothetical[b]	300	5	2310	3.4%	3.7%	Hendrickson *et al.* (1989)
Streptavidin	252	2	1850	2.5%	2.3%	Hendrickson *et al.* (1989)
Crambin	46	6 S	350	1.5%		Hendrickson and Teeter (1981)

The Bijvoet diffraction ratio is computed from $<|\Delta F_{\Delta h}|> = q \cdot 2f''$ and the dispersive diffraction ratio is $<|\Delta F_{\pm\Delta\lambda}|> = q \cdot |f'(\lambda_i) - f'(\lambda_j)|$ where $q = \sqrt{N_A/2N_P}/Z_{eff}$ (Hendrickson *et al.*, 1985). Here N_P, the number of non-hydrogen atoms in the protein molecule, is calculated assuming 7.7 atoms per amino acid residue and Z_{eff}, the effective normal atomic scattering at zero scattering angle, is 6.7 electrons. N_A is the number of anomalous scatterers per protein molecule.

[a]These hypothetical molecules have methionine concentrations at 1/5, 1 and 2 times average, and diffraction ratios are calculated assuming measurements made at wavelengths corresponding to the extrema of the average spectrum of *E.coli* thioredoxin as given in Table I. A remote wavelength of 0.90 Å ($f' = -1.62$) is assumed for the dispersive difference.

[b]This hypothetical model assumes methionine concentration at the average and scattering factors with low energy resolution as in the streptavidin experiment.

(Graves *et al.*, 1990), to IL-2 (M.H.Hatada and B.J.Graves, personal communication) and to ribonuclease H (W.Yang, unpublished results). Clearly, the selenomethionyl proteins are sufficiently stable to permit the necessary diffraction measurements.

The scattering properties of selenium are also very favourable. The energy level of the selenium K-shell (1s) orbital corresponds to an X-ray wavelength (0.98 Å) that is readily accessible with synchrotron radiation. This also has advantages of limited radiation damage and sample absorption. The relative abundance of methionine residues in proteins is such that adequate diffraction signals can be expected in typical cases (Table III). Of course, an atom can only contribute coherently to the diffraction if it is well ordered, and thus atomic mobility is a concern. Selenomethionine is a good label in this regard. A survey of crystallographic B values (atomic mobility parameters) reported for 189 protein structures in the Protein Data Bank (Bernstein *et al.*, 1977) shows that on average the atomic mobility at the Sγ position of a methionine is just 4% above the average for all side-chain atoms in the molecule. For comparison, mobilities at the Cγ positions in leucine average 16% less, and those at Cγ in glutamine, arginine and lysine average 33%, 16% and 42% higher, respectively (J.R.Horton, unpublished results). A complication can arise in the diffraction analysis of large proteins where the number of selenium sites may be so large that this substructure presents a difficult problem. However, a simulation based on the structure of citrate synthase (Wiegand *et al.*, 1984), which with 30 had the most methionines per asymmetric unit among proteins in a particular release from the Protein Data Bank, shows that this selenium substructure could readily be solved by direct methods with the program MULTAN (Germain *et al.*, 1970) even when 20% random error was applied (J.R.Horton, unpublished results).

Possible complications in selenomethionyl protein expression

While met$^-$ strains of *E.coli* can clearly be grown on selenomethionine, it is also evident that care must be taken to achieve complete replacement. The major complication arises from the difficulty in bringing cells out of stationary phase in the presence of selenomethionine, whereas methionine brought forward with the inoculum for log phase growth may be preferentially incorporated. Thus, when we produced thioredoxins with excess methionine present in the starter culture, only 90% selenomethionine incorporation was obtained. Similar complications also limited the production of selenomethionyl IL-1α to an 86% level (Graves *et al.*, 1990). On the other hand, we have produced recombinant *E.coli* ribonuclease H with 100% selenomethionine incorporation by using a defined selenomethionyl starter medium supplemented by 5% Luria-Bertani medium (W.Yang, unpublished results) while in other experiments incorporation was limited to ~90%.

The vitality of the DL41 strain, which derives directly in three steps from the ancestral K12 stock, is an advantage for selenomethionine replacement. Other auxotrophs tested did not tolerate selenomethionine as well. However, the lambda P_L expression system is probably a disadvantage. The potential temperature sensitivity of selenomethionyl proteins can complicate induction of the recombinant plasmid by temperature jump.

Another potential complication is that, as with selenomethionine itself (Shepherd and Huber, 1969), the solubilities of selenomethionyl proteins are somewhat reduced in comparison with their sulfurous counterparts. Although this has not yet prevented us from obtaining crystals of the selenomethionyl products under conditions similar to those used for the natural proteins, this might not always be the case. A more serious complication might arise from susceptibility of selenomethionine to oxidation. We have noted a loss of birefringence and diffraction (despite retention of clean morphology) from crystals of *E.coli* thioredoxin kept for two months in a stabilizing medium. The cause of this change has not been investigated but storage in an oxygen-free environment is obviously indicated.

Finally, although procedures for the production of selenomethionyl proteins in *E.coli* are now well developed, little has been done in eukaryotic cells. As discussed above, mammals are naturally auxotrophic for methionine. However, the greater complexity of mammalian cells tends to make them more sensitive to selenomethionine. Nevertheless, we have achieved 87% incorporation of

selenomethionine into a recombinant soluble CD4 (Deen *et al.*, 1988) produced in CHO cells (P.D.Kwong, unpublished results). Since yeasts contain pathways for production of methionine from cysteine as well as the reverse pathways as found in mammals, appropriate mutants will probably be required for the expression of selenomethionyl proteins in yeast cells.

Selenocysteine as a prospect for MAD phasing

Those few proteins that naturally contain selenocysteine are obvious candidates for analysis by MAD phasing. For example, in light of results in Table III, there is little doubt that the structure of glutathionine peroxidase with one selenocysteine in each 182 residue chain (Epp *et al.*, 1983) could have been solved in this way. Unfortunately, the failure of *E.coli* grown on selenate to incorporate selenocysteine into β-galactosidase (Huber and Criddle, 1967) makes a general replacement scheme such as used for selenomethionine seem unlikely. On the other hand, the recent discovery of utilization of the UGA termination codon in the biosynthesis of selenocysteinyl proteins (Chambers *et al.*, 1986; Zinoni *et al.*, 1986) suggests a general approach through the engineering of selenocysteinyl mutants. This, however, is complicated by the apparent involvement of the nucleotide context of the UGA codon in the signal of selenocysteine incorporation. An alternative possibility suggested by Jeffrey Miller (personal communication) would be to construct and use an amber suppressor gene (Normanly *et al.*, 1986) starting from the *E.coli selC* gene which codes for the special tRNA species leading to selenocysteine insertion (Leinfelder *et al.*, 1988). Both plans suffer from possible adverse effects of the susceptibility of selenocysteine to oxidation. Epp *et al.* (1983) found that the selenocysteine of glutathione peroxidase is oxidized to seleninic acid. It might be possible, however, to capitalize on this proclivity toward oxidation to form especially stable diselenide selenosulfide analogs of disulfide bridges.

Materials and methods

Methionine auxotroph transduction

The DL41 strain was constructed by phage P1 transduction of mal E: Tn*10* from strain TST1 (CGSC strain no. 6137) into MG1655 (Guyer *et al.*, 1981) followed by transduction of *metA* from AB1932 (Howard-Flanders *et al.*, 1966) by selection for maltose utilization and screening for methionine auxotrophy.

Thioredoxin overproduction and purification

A starter culture of the DL41 strain containing the pDL59 expression plasmid (LeMaster and Richards, 1988) was grown overnight at 33°C in a defined medium that contains 40–50 mg/l methionine. This starter culture was then diluted 1/200 (v/v) into prewarmed medium devoid of methionine but containing 40–50 mg/l D,L selenomethionine (Sigma). This leads to only a 0.5% residue of sulfurous methionine in the growth medium; it is critical to minimize this methionine content. Cells in this medium were grown at 33°C to OD_{600} of 1.0, at which point transcription from the plasmid was induced by abrupt shift of temperature to 43°C. Expression was allowed to continue for 2.5 h.

Purification of the *E.coli* and T4 thioredoxins expressed in these fermentations was carried out as previously described (LeMaster and Richards, 1988). In brief, the debris from sonicated cells was pelleted and the supernatant was treated by streptomycin sulfate and ammonium sulfate fractionation. The thioredoxin-containing fraction was then resuspended, further fractionated on a sizing column, and then separated on a DE52 column into the two constituent thioredoxins. Finally, the samples were lyophilized after dialysis against 25 mM NH_4HCO_3.

Amino acid analysis

Analyses for amino acid composition were performed on a Beckman 6300A analyzer after hydrolysis *in vacuo* with 6 N HCl at 110°C for 24h.

Crystallization

Crystallizations were carried out with the hanging-drop vapor diffusion technique in 24-well Linbro cell culture plates. All chemicals were obtained from Sigma Chemical Co., St Louis, MO unless otherwise noted.

T4 thioredoxin

Lyophilized natural and selenomethionyl T4 thioredoxins were dissolved in buffer containing 40 mM bis-tris propane, pH 6.8, to a concentration of ~13 mg/ml. Droplets (3 μl) of this saturated solution were placed on siliconized glass coverslips which were then inverted over wells containing a 0.5 ml solution of 12–28% ethanol, the bis-tris propane buffer and 2 mM cadmium acetate (Aldrich) and sealed with silicone vacuum grease. The plates were stored at 4°C for 5–7 days at which time 1 μl of the well solution was added to the drops. Clusters of small crystals appeared to grow from amorphous material in some of the 16–18% ethanol setups ~5 days later.

E.coli thioredoxin

Starting droplets for crystallization of the *E.coli* proteins were constituted from 5 μl of protein solution (~10 mg/ml), 5 μl of a pH 4.2 buffer including 10 mM sodium acetate and 2 mM cupric acetate, and 1 μl of 40% (400 mg/100 ml) polyethylene glycol 3350 (PEG 3350). Such droplets formed on siliconized coverslips were inverted and sealed over 1 ml reservoirs containing 12–24% PEG 3350, 5 mM acetate buffer (pH 4.2), and 1 mM cupric acetate. These experiments were carried out in an incubator held at 20°C. Crystals of selenomethionyl *E.coli* thioredoxin first appeared after 4–5 days whereas those of the native protein appeared in 10–14 days. The best crystals grew over wells of 16–20% PEG without appreciable distinction between the two proteins. However, crystals only grew in some setups which suggested to us that nucleation was a rate limiting process. The yield of usable crystals was improved by microseeding. Grains of washed and dried sand (Baxter, Edison, NJ) that were added to some droplets also appeared to increase the yield of crystals.

The crystals used to determine unit cell dimensions (Table I) and at SSRL for X-ray absorption spectroscopy (Figure 4) and synchrotron data collection were grown from material produced as described above from protein solutions of 11 mg/ml over wells of 16% PEG. The crystals and diffraction patterns shown in Figure 3 were obtained from protein produced in subsequent fermentations. These crystallizations were carried out over 20% PEG wells starting from stock solutions of 13.9 and 8.6 mg/ml for the natural and selenomethionyl proteins respectively. The selenomethionyl thioredoxin in this case proved to have only 90% substitution in this instance due to an excess of methionine in the starter culture.

Crystal characterization

Crystal growth and morphology were monitored by examination with a Zeiss stereomicroscope. Crystals were mounted in thin-walled glass capillaries for diffraction analysis. Symmetry properties were determined by photography on a Huber precession camera. Unit cell dimensions were determined from ~20 reflections centered on a Rigaku AFC-5 diffractometer. X-rays used in these experiments were emitted from the copper anode of a Rigaku RU-200 generator.

X-ray absorption spectroscopy

X-ray absorption spectra were measured by scintillation-counter detected X-ray fluorescence using the area detector system (Phizackerley *et al.*, 1986) at the Stanford Synchrotron Radiation Laboratory. The method of measurement and the procedures for reduction of the normalized fluorescence spectrum to scattering factor curves were those described in Hendrickson *et al.* (1988).

Acknowledgements

We thank Ray Salemme for suggesting thioredoxin as a suitable test system for selenomethionyl protein production, Paul Phizackerley and Ethan Merritt for help with synchrotron experiments at Stanford, Bob Batterman for measuring the absorption spectrum from selenomethionine at the Cornell synchrotron, Marcos Hatada and Brad Graves for sharing results on selenomethionyl interleukins, Mary Ann Gawinowitz for amino acid analyses, Jeffrey Miller for discussions about selenocysteine, and our colleagues Peter Kwong, Tom Peat and Wei Yang for unpublished results and helpful discussions. This work was supported in part by a grant (GM34102) from the National Institutes of Health, and the synchrotron facilities used at Stanford are supported by the Department of Energy and by the National Institutes of Health.

References

Bernstein,F.C., Koetzle,T.F., Williams,G.J.B., Meyer,E.F., Brice, M.D., Rogers,J.R., Kennard,O., Shimanouchi,T. and Tasumi,M. (1977) *J. Mol. Biol.*, **112**, 535–542.

Bremer,J. and Natori,Y. (1960) *Biochim. Biophys. Acta*, **44**, 367–370.

Chambers,I., Frampton,J., Goldfarb,P., Affara,N., McBain,W. and Harrison,P.R. (1986) *EMBO J.*, **5**, 1221–1227.

Cowie,D.B. and Cohen,G.N. (1957) *Biochim. Biophys. Acta.*, **26**, 252–261.

Cromer,D.T. (1983) *J. Appl. Crystallogr.*, **16**, 437.

Dayhoff,M.O. (1978) *Atlas of Protein Sequence and Structure*. National Biomedical Research Foundation, Washington, D.C., Vol. 5, Suppl. 3, p. 363.

Deen, K. C., McDougal, J.S., Inacker, R., Folena-Wasserman, G., Arthos,J., Rosenberg, J., Maddon, P. J., Axel, R. and Sweet,R.W. (1988) *Nature*, **331**, 82–84.

Epp,O., Ladenstein,R. and Wendel,A. (1983) *Eur. J. Biochem.*, **133**, 51–69.

Frank,P., Licht,A., Tullius,T.D., Hodgson,K.O. and Pecht,I. (1985) *J. Biol. Chem.*, **260**, 5518–5525.

Germain,G., Main,P. and Woolfson,M.M. (1970) *Acta Crystallogr. B*, **26**, 274–285.

Graves,B.J., Hatada,M.H., Hendrickson,W.A., Miller,J.K., Madison,V.S. and Satow,Y. (1990) *Biochemistry*, in press.

Guss,J.M., Merritt,E.A., Phizackerley,R.P., Hedman,B., Murata,M., Hodgson,K.O. and Freeman,H.C. (1988) *Science*, **241**, 806–811.

Guyer,M., Reed,R.R., Steitz,J.A. and Low,K.B. (1981) *Cold Spring Harbor Symp. Quant. Biol.*, **45**, 135–140.

Hahn,G.A. and Brown,J.W. (1967) *Biochim. Biophys. Acta*, **146**, 264–271.

Hendrickson,W.A. (1985) *Trans. Am. Crystallogr. Assoc.*, **21**, 11–21.

Hendrickson,W.A. and Teeter,M.M. (1981) *Nature*, **290**, 107–113.

Hendrickson,W.A., Smith,J.L. and Sheriff,S. (1985) *Methods Enzymol.*, **115**, 41–55.

Hendrickson,W.A., Smith,J.L., Phizackerley,R.P. and Merritt,E.A. (1988) *Proteins*, **4**, 77–88.

Hendrickson,W.A., Pähler,A., Smith,J.L., Satow,Y., Merritt,E.A. and Phizackerley,R.P. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 2190–2194.

Holmgren,A. (1968) *Eur. J. Biochem.*, **6**, 475–484.

Holmgren,A., Söderberg,B.-O., Eklund,H. and Brändén,C.-I. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 2305–2309.

Howard-Flanders,P., Boyce,R.P. and Theriot,L. (1966) *Genetics*, **53**, 1119–1136.

Huber,R.E. and Criddle,R.S. (1967) *Biochim. Biophys. Acta*, **141**, 587–599.

Karle,J. (1980) *Int. J. Quant. Chem. Symp.*,**7**, 357–367.

Katti,S.K., LeMaster,D.M. and Eklund,H. (1990) *J. Mol. Biol.*, in press.

Krishna Murthy,H.M., Hendrickson,W.A., Orme-Johnson,W.H., Merritt,E.A. and Phizackerley,R.P. (1988) *J. Biol. Chem.*, **263**, 18430–18436.

Leinfelder,W., Zehelein,E., Mandrand-Berthelot,M.-A. and Böck,A. (1988) *Nature*, **331**, 723–725.

LeMaster,D.M. (1986) *J. Virol.*, **59**, 759–760.

LeMaster,D.M. and Richards,F.M. (1985) *Biochemistry*, **24**, 7263–7268.

LeMaster,D.M. and Richards,F.M. (1988) *Biochemistry*, **27**, 142–150.

Lunn,C.A., Kathju,S., Wallace,B.J. Kushner,S.R. and Pigiet,V. (1984) *J. Biol. Chem.*, **259**, 10469–10474.

McConnell,K.P. and Hoffman,J.L. (1972) *FEBS Lett.*, **24**, 60–62.

Mudd,S.H. and Cantoni,G.L. (1957) *Nature*, **180**, 1052.

Normanly,J., Masson,J.-M., Kleina,L.G., Abelson,J. and Miller,J.H. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 6548–6552.

Odom,J.D. (1983) *Structure and Bonding*, **54**, 1–26.

Phizackerley,R.P., Cork,C.W. and Merritt,E.A. (1986) *Nucl. Instrum. Methods A*, **246**, 579–595.

Rajeswaran,M. and Parthasarathy,R. (1984) *Acta Crystallogr. C*, **40**, 647–650.

Shepherd,L. and Huber,R.E. (1969) *Can. J. Biochem.*, **47**, 877–881.

Sjöberg,B.-M. and Holmgren,A. (1972) *J. Biol. Chem.*, **247**, 8063–8068.

Sjöberg,B.-M. and Söderberg,B.-O. (1976) *J. Mol. Biol.*, **100**, 415–419.

Söderberg,B.-O., Sjöberg,B.-M., Sonnerstam,U. and Brändén,C.-I. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 5827–5830.

Stadtman,T.C. (1987) *FASEB J.*, **1**, 375–379.

Templeton,L.K. and Templeton,D.H. (1988) *Acta Crystallogr. A*, **44**, 1045–1051.

Wiegand,G., Remington,S., Deisenhofer,J. and Huber,R. (1984) *J. Mol. Biol.*, **174**, 205–219.

Zinoni,F., Birkman,A., Stadtman,T.C. and Böck,A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 4650–4654.

Received on January 5, 1990